Predictive profiling and its legal limits

Lammerant, Hans; de Hert, Paul

Published in:
Exploring the boundaries of big data

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

6 PREDICTIVE PROFILING AND ITS LEGAL LIMITS: EFFECTIVENESS GONE FOREVER?

Paul De Hert & Hans Lammerant

Societal activities are increasingly mediated by digital technology, leading to new forms of visibility left by their digital traces. The spread of digital technology also gives rise to new forms of cognition, made possible by data analysis techniques that allow sense-making and the steering of governance. 'Big Data' is the new term pointing to the growing availability of data and new data-driven practices. Profiling is one of these new forms of cognition. Although it is not new, profiling is flourishing in the Big Data environment and used on a much wider scale than ever before.

We examine predictive group profiling in the Big Data context as an instrument of governmental control and regulation. We first define profiling by drawing some useful distinctions (section 6.1). We then discuss examples of predictive group profiling from policing (such as parole prediction methods taken from the US) and combatting fraud (the ICOV and SyRI systems in the Netherlands) (section 6.2). Three potential risks of profiling – the negative impact on privacy; social sorting and discrimination; and opaque decision-making – are discussed in section 6.3.

We then turn to the legal framework. Is profiling by governmental agencies adequately framed? Are existing legal checks and balances sufficient to safeguard civil liberties?[1] We discuss the relationship between profiling and the right to privacy (section 6.4) and between profiling and the prohibition on discrimination (section 6.5). The jurisprudence on the right to privacy clearly sets limits to the use of automated and predictive profiling. Profiling and data screening which interfere without distinction with the privacy of large parts of the population are disproportional. Applications need to have some link to concrete fact to be legitimate. An additional role is played by the prohibition of discrimination, which requires strengthening through the development of audit tools and discrimination-aware algorithms. We then discuss current safeguards in Dutch administrative, criminal procedure and data protection law (section 6.6), and witness a trend of weakening safeguards at the very moment when they should be applied with even more rigor. In our conclusion we point to the tension between profiling and legal safeguards. These safeguards remain important and need to be overhauled to make them effective again.

[1] Van der Sloot, B and others, exploring the boundaries of big data 2016 Amst UP

6.1 PROFILING AND BIG DATA: CONCEPTS, DISTINCTIONS AND EVOLUTIONS

Human and automated profiling

It is useful to remind ourselves that profiling is not new and can be done without modern automated and statistical techniques ('human profiling'). Practices of ethnic profiling by police during stops and searches on the street are an example of human profiling, albeit a problematic one.[2]

In general, a profile is a set of characteristics, features and attributes with which a person or a group can be discerned from another person or group. What interests us here is a specific process of differentiating based on the automated processing of data. The form of profiling we have in mind is based on the use of 'Knowledge Discovery in Databases' (KDD), better known as data mining (although this is only the analysis step in the process) (Fayyad et al. 1996). The purpose of KDD is to find useful patterns in data, which can be gathered from different sources. The first stages of the process entail selecting and gathering data and preparing it for analysis. In the actual data mining, data is analysed with the use of algorithms in order to discern patterns. While these patterns reflect correlations in the data, they are no proof of causal relations. A pattern can be an indication of a relevant underlying causal process, but it can also be the result of uninteresting processes or noise. Therefore the final step consists in evaluating these patterns for their relevance. From these selected patterns a profile can be derived. This profile does not consist of 'raw' data or observations, but is a mathematical model of the phenomenon or a reference to the group to discern.

Data mining makes use of new statistical techniques driven more by data than by theory. Traditionally, a hypothesis is formulated first, which is then verified with the data. The emphasis is on clarifying assumptions and using statistical methods to differentiate between significant correlations and correlations that are spurious or cannot be guaranteed to differ from chance. The new data-driven approach, for which the term data mining has become common usage, begins with the data and searches for patterns. Selection is done later as part of the interpretation. This allows discovering unforeseen relations in the data, but also introduces the risk of using correlations and models without understanding the actual process that produced them. Many data mining algorithms are opaque: it is difficult or impossible to determine how the resulting model was built and which correlations were taken into account. We also need to remember that the results are purely probabilistic relations which need verification. In other words, they deliver no proof but only indications for where to look. Whereas the traditional approach uses statistics to determine or measure the veracity or plausibility of a hypothesis, it becomes a discovery method in itself in the data mining approach.

Creating and using a profile: three groups of affected people

What we have described so far is the creation of a profile. It has to be distinguished from the use of profiles in decision-making processes, which consists of applying profiles to datasets and checking which persons, objects or phenomena conform to the profiles, and making decisions based on the result. Profiles can be used to identify people, to attribute specific risks to them, and to act upon them in specific ways. Custers identifies four possible uses of profiling (Custers 2014). First, as a selection instrument to decide which persons or groups deserve more attention to guide controlling efforts.[3] Second, as an instrument in decision-making, where decisions are made based on the profile without further investigation. The space for such use is limited as the data protection framework forbids automated decision-making without further human intervention.[4] A third use of profiling is as a detection instrument, to detect if certain rules have been violated, not who has violated them. Fourth, profiling can be used to evaluate practices and interventions.

The distinction between the creation of a profile and its use is crucial for our legal analysis, as its creation and use happen at different moments and under different circumstances, and make use of different sets of data. To make a profile, data is used from a wide range of people, including citizens who are not suspected of violating any laws. The application of a profile uses only the data of the persons being checked, which can be a large set or just one person. This distinction notwithstanding, profiles are linked to their use. Their envisaged use will define what must be differentiated and thus provide the criteria to discern relevant patterns during the creation of a profile. The purpose also matters when evaluating the legitimacy of profiling.

Three groups are affected by the use of profiling: persons whose data is used to create the profile, persons to whom the profile refers, and persons who are subjected to decision-making based on the profile. These three groups often overlap but are not by definition the same.[5]

Personal and group profiling

Group profiles must be distinguished from personal profiles. A personal profile concerns an individual subject. Through a set of features, this subject can be identified and targeted. An example is device fingerprinting, where a device is recognized by specific technical features such as installed software or an IP number. Face recognition and other forms of biometric profiling also create personal profiles. A group profile concerns a category of people. The data mining process uncovers a range of correlated attributes and links these in a specific pattern that constitutes the category or group profile. While this can be an existing group, in many cases the category is established through the process of profiling itself and has no existence before the creation of the profile (Hildebrandt 2008).

We need to differentiate between distributive and non-distributive group profiles. When all members share the features, the group profile is distributive; the profile gives an exact representation of the features of individual group members. In this case the group profile can be applied to group members as if it is a personal profile. In contrast, a non-distributive profile only represents a statistical relation. Group members do not share all features, but have most of them. Membership represents a correlation between the set of attributes of a member with the attributes of other group members. The difference becomes clear when subjects who accord with a non-distributive profile are treated as if they match a distributive one. An average characteristic or behaviour as represented by the profile is then considered as a characteristic of each group member. The result is stereotyping, which can lead to discrimination (as in the aforementioned ethnic profiling) or normalizing effects on group members (Hildebrandt 2005; Vedder 1999).

Predictive profiling entails the use of profiling to generate predictions. The profile, extracted from data on past behaviour or known cases, is used to infer current or future behaviour or the state of unknown cases. As such, the profile is a probabilistic model. It does not represent actual behaviour or a real state of affairs, but a prediction of this behaviour or state of affairs. Again, treating this probabilistic representation as an exact one can be problematic, as it assumes that nothing changes.

The dual impact of Big Data on profiling

Automated profiling based on data mining is not yet by definition an application of Big Data. Although Big Data is not a precisely defined concept and is often used as a buzzword, it reflects a change in techniques to collect and analyse data. The term originates from developments in database technology, where data became too voluminous to store or analyse on a single computer. Big Data is often characterised by its 3 V's definition (volume, velocity, variety). It has to deal with much larger volumes of data, collected or produced at much higher velocities, and consisting of a much wider variety of, often unstructured, data (Kitchin 2013).

Legally it is irrelevant whether profiling is part of a Big Data application or not. But Big Data magnifies the impact of profiling and puts greater stress on the checks and balances in the legal framework. One important change is that Big Data strengthens the evolution towards a data-driven epistemology, from checking hypotheses to exploring data for correlations (Kitchin 2014). While this was already signalled with the advent of data mining, Big Data creates a new environment in which these techniques become much more powerful.

Another change can be seen in data aggregation and collection methods. Data is aggregated through linking existing data sources and making them inter-operable. Data collection also becomes more sensor-driven, partly through the mediation of digital technology. New types of sensors are entering widespread use, such as

ANPR cameras, or are slowly becoming a technological reality, like smart cameras with face recognition. These changes – from once anonymous train or parking tickets to the more permanent and individualised visibility of behaviour through the use of OV cards and SMS parking – are creating new visibilities. We now look at some concrete examples of profiling in the US and the Netherlands.

6.2 PROFILING IN THE PUBLIC SECTOR: EXAMPLES FROM POLICING AND ADMINISTRATIVE ANTI-FRAUD POLICIES

The advent of profiling as a governmental technique of control is not only due to new technological possibilities, but also fits changing approaches to management and security. Julia Black has written authoritatively on the 'new public risk management' in public administration, where assessments and technologies are used to develop decision-making frameworks to prioritize regulatory activities and deploy resources (Black 2005). A range of other authors, mainly in criminology and police studies, have addressed changes in security policies. Traditional post-crime policing approaches, directed at investigation and the punishment of crime, have been deemed unable to deter crime and have been exchanged for more preventive, pro-active strategies. One of these is intelligence-led policing; its main characteristic is "its insistence to build up intelligence through all kinds of data collection strategies" (Van Brakel and De Hert 2011).

These trends in management and policing interact with technological developments, allowing for much wider data collection, aggregation and analysis. Intelligence-led policing is very open to the growing use of surveillance techniques and leads to a widening net of observation and data collection. Similarly, intelligence-led policing is open to new analytical uses of this data such as predictive and automated profiling. The preventive approach is not only seen in policing strategies, but also in criminal and punitive policy. Again we see the interaction with new technologies, such as the use of profiling techniques for parole decisions in the US (Van Brakel and De Hert 2011).

The interaction of Big Data with intelligence-led policing is equally visible in the growing effort to make data available for law enforcement and security purposes. On the EU level, there has been a steady build-up of European-wide information systems like the SIS II, VIS and Eurodac. The exchange of information between police services has become the norm with the introduction of the principle of availability. Specific legal frameworks concerning information collection and exchange have been set up, e.g. on terrorist financing, the retention of telecommunications data (now annulled by the ECJ), and the ongoing discussion on Passenger Name Records. We now take a closer look at some examples in the US and the Netherlands, first from the security field.

In the US, automated profiling techniques for policing have grown popular under the label of 'predictive policing'. Predictive policing builds on older statistical approaches using crime statistics, but has developed using new data mining techniques and more sources of data. A RAND study on predictive policing gives a good overview of approaches and cases. It found four main methods of predictive profiling in policing: (1) methods for predicting crimes;[6] (2) methods for predicting offenders;[7] (3) methods for predicting perpetrators' identities;[8] and (4) methods for predicting victims of crimes[9] (Perry et al. 2013).

The most widespread forms of predictive policing are type 1 (predicting crimes) and 4 (predicting victims). They aim to uncover crime hotspots, in terms of time and place, where future criminal activity can be expected. Crime statistics are combined with GIS and data sources like traffic data. This predictive profiling of hotspots is used to guide police patrolling and observation. Another example of hotspot profiling reveals which types of shops and branches are most vulnerable to robberies.[10] These methods have been reported to deliver positive results and are becoming general use in the US (Perry et al. 2013).

Although considered much less developed, profiling methods are also entering into use for the criminal profiling of offenders (Perry et al. 2013: 81). The methods for predicting the risk that an individual will offend in the future are mostly based on risk assessments methods. The main challenges for these methods are inter-rater reliability (agreement between raters, a measure of the reliability of this data) and misspecification of the predictive model (Perry et al. 2013). These challenges notwithstanding, parole prediction methods are used in a majority of US states (Harcourt 2005). Previously based on risk assessment methods, they are also now based on software (McCaney 2013). Florida, for example, uses predictive methods to assign juveniles to rehabilitation programs (Perry et al. 2013). Another example, as advanced as it is controversial, is the Chicago police pilot program to create a 'Heat List' – "a rank-order list of potential victims and subjects with the greatest propensity for violence." The Heat List is "generated based on empirical data compared with known associates of the identified person" – that is, on social network analysis (Chicago Police Department 2013). The program is inspired by the work of sociologist Andrew Papachristos, which showed that the majority of homicides occur in a relatively small network within the population. People on the Heat List are contacted and warned not to commit an offence, sometimes to their surprise if they have previously never been in contact with the police (The Verge 2014). Methods to create profiles of likely perpetrators of past crimes are generally not based on automated modelling but on crime analysts combining and querying datasets. They use larger and more diverse sets of data but the modelling itself is not done by algorithms but by human intelligence. Examples include geographical profiling, or trying to locate perpetrators based on the pattern of crimes, and modus operandi similarity analysis (Perry et al. 2013).

Similar methods can be used outside the security context and are now used by the public sector to combat tax and social fraud. A GAO report found the US federal government using a wide range of data mining applications, many but not all using personal data. Typical applications aim to detect fraud or abuse or to improve service or performance (GAO 2004).

A Dutch example is the *Infobox Crimineel en Onverklaarbaar Vermogen* (iCOV), a cooperative structure between the tax and customs authorities, police and public prosecutor to map criminal and unexplained finances, money laundering and fraud constructions and to recover public financial claims. Its tasks include developing indicators and group profiles, which can be based on personal data (Convenant iCOV 2013).

Another Dutch example to combat social fraud is the linking of a wide range of personal information through the SyRI (System Risk Indication) instrument. This system is based on the *Wet structuur uitvoeringsorganisatie werk en inkomen* (SUWI), or the law covering the organisational set-up of government tasks in social insurance and employment. The SyRI system combats social fraud and the abuse of public money, with a range of public authorities pooling their data resources and screening them using risk models. The range of information that can be used is very broad, and includes data on labour, taxes, social security, property registers and debt. Specific sources are selected and a risk model that defines the profile of potential fraud is developed for each project. When approved by the minister, the data goes encrypted to an analysing unit. There the profile is applied to the data and people linked with a potential fraud pattern are flagged. Following analysis, partners receive risk notifications on flagged persons for further action. These are also kept in a register, where people can inquire if they are included.

The SyRI system has a well-developed organisational framework and is said to comply with existing data protection obligations. In terms of proportionality, it is striking that the only limiting factor on the data included for analysis is the risk model or profile of potential fraud. The model has to be developed before the screening of data begins; the system is thus not used without precautions and excludes the unlimited screening of data. However, no suspicion or indications are needed to start a project; nor are factual grounds. The risk models contain general assumptions about indications of fraud, without requiring a concrete problematic based on facts. This has been criticised by the Dutch Data Protection Authority, which insists on preliminary 'evidence' to start the system and the application of the 'Select before you collect' principle. This implies that the link between indicators and possible fraud must be legitimate, and that only those types of data tied to this indicator can be selected. It also implies selecting the persons whose data will be used beforehand and to legitimize this choice (CBP 2014).

6.3 POTENTIAL RISKS OF PROFILING: PRIVACY, SOCIAL SORTING AND DISCRIMINATION, OPAQUE DECISION-MAKING

Profiling techniques can increase efficiency as they allow for more precise targeting and the economizing of efforts in a whole range of areas. Custers describes how prioritizing inspections based on risk models of illegally splitting living space made inspections more effective (Custers 2014). Similarly, guiding policing efforts through predictive policing in the US is generally considered successful (Perry et al. 2013). But it is not without risks, of which we want to highlight three: privacy, social sorting and discrimination, and opaque decision-making.

Firstly, there is a heightened risk of intrusive interferences to privacy due to the heightened surveillance. Profiling is in itself highly intrusive to a wide range of people, especially if it is used routinely. The technical possibilities encourage making a growing amount of data sources available, in terms of both linking existing sources of data and the implementation of new systems soliciting behavioural data. The result is a widening net of data collection and surveillance, which can have a heavy psychological impact. This impact comes not from the data collection itself but from its integration and transformation into knowledge about people (Hildebrandt 2008a). Hildebrandt points to privacy as a public good and its relationship to autonomy and human agency. Profiling can result in information asymmetries threatening such autonomy (Hildebrandt 2008b). This risk affects all people whose data is used to build the profile as well as people to whose data the profile is applied.

Secondly, there is the risk associated with social sorting, or sorting people into categories assigning worth or risk, and stereotyping. The formation of a suspicion, which can legitimize further public intervention, is no longer based on specific facts but on people's characteristics. It thereby threatens the principle of equality and the presumption of innocence and can lead to discrimination (Schermer 2013; Custers 2014; Gandy 2009). Such discrimination can take several forms. It can lead to a high control burden for specific groups who often match the profiles. It can also be hidden discrimination due to the opacity of data mining algorithms.

Thirdly, there are the risks related to opaque decision-making, which plays out on two distinct levels: the application of a profile and the decision-making based on its results. The opacity of data mining techniques can create problems through lack of insight into the veracity of the model. Some data mining techniques result in profiles where it is very unclear which data were most important and which data leads to the flagging of a person. Negative influences on reliability can also result from missing data, bias in the data, or wrong data. This can lead to an accountability risk, when too much responsibility is placed on these techniques. This problem may seem limited, as predictive profiling is mostly used to guide investigative

measures and further decisions will be based on their findings. But when predictive profiling is used to legitimize intrusive preventative measures, e.g. based on risk scores, not just the creation of a profile but the whole decision-making process becomes opaque.

Citron (2008) warns how procedural safeguards are imperilled by automation in what she calls the 'automated administrative state'. Involving more humans in the decision-making process will not do, Citron argues, since in practice there is little difference between automated decision-making and non-automated decision-making with human intervention before the decision. People, including state officials, are often fooled by technology or do not possess the ability or information to properly assess the computer-generated suggestion. Too much trust in the results provided by the computer leads to an 'automation bias'. Coding can result in a hidden form of rule-making through programmers' mistakes or interpretations. The construction of a profile can be seen as a similar form of hidden rule-making. Citron shows that procedural safeguards need a major overhaul to remain effective.

6.4 PROFILING AND THE RIGHT TO PRIVACY

The general framework and the problematic proportionality test

Profiling clearly enters the ambit of the fundamental right to privacy. Article 8(1) of the European Convention on Human Rights (ECHR) defines the right to privacy as "the right to respect for his private and family life, his home and his correspondence." Article 8(2) defines when interfering with this right is legitimate.

The first question is if profiling interferes with the right to privacy. Profiling is not in itself a form of data collection, but uses data collected earlier, which we assume took place legally. It means that the data is now in the possession of public authorities, sometimes in the possession of other administrations and collected for other purposes. Whereas data in the US is no longer considered private once it is held by other parties or is public (Solove 2004), the European Court of Human Rights (ECtHR) does not limit private life to the intimate sphere. Even in the public sphere a person's privacy still has a certain protection. The ECtHR pointed out in *P.G. & J.H. v. United Kingdom* that there is "a zone of interaction of a person with others, even in a public context, which may fall within the scope of 'private life'". It also stated that:

> "Private-life considerations may arise, however, once any systematic or permanent record comes into existence of such material from the public domain. It is for this reason that files gathered by security services on a particular individual fall within the scope of Article 8, even where the information has not been gathered by any intrusive or covert method."

This clarifies that the use of data linked to persons is an interference to their private life and needs to be legitimized. Public authorities therefore have a duty to consider the privacy rights of individuals when deciding on the further use of such data. Similar considerations were raised in other case law on the storage (*Leander v. Sweden, Rotaru v. Romania, Amann v. Switzerland*), further retention (*S. and Marper v. United Kingdom*), and diffusion of personal data (*Peck v. United Kingdom, Z. v. Finland, Weber and Saravia v. Germany*). We can conclude that profiling activities by public authorities are to be considered interferences to private life, also when they use data on behaviour in the public sphere (such as mobility data) or data that was earlier handed to public authorities (e.g. in the context of social security or taxation). This concerns both the creation of profiles and the application of profiles to specific persons.

What makes an interference to privacy legitimate? There are three requirements. First, the interference has to be in accordance with the law. There has to be some ground in domestic law for the measure. The ECtHR also looks at the 'quality of the law' or the substantive content, which has "to provide safeguards against arbitrariness" (*P.G. & J.H. v. United Kingdom*). In practice this rather technical requirement has turned out to be an important check on both human and automated profiling. In *Gillan and Quinton* v. *United Kingdom*, the ECtHR held that stop and search powers by law enforcement in the UK violated privacy, because they are "not in accordance with law". Under sections 44-47 of the *Terrorism Act 2000*, police in the UK gained the power to stop and search people without any requirement to first form a reasonable suspicion of unlawful behaviour. The Court held that the extraordinary breadth of power given to the police under the Act lacked appropriate legal safeguards capable of protecting individuals against arbitrary interference. The Court also acknowledged the risks of the discriminatory use of powers against black and Asian persons. Statistics accepted by the Court showed that black and Asian persons were disproportionately affected by police powers in the UK.

The same strict line was also followed in a case concerning digital surveillance, a case that led to *Liberty v. United Kingdom*. Here the ECtHR held that a system of mass surveillance operated by the UK government to spy on all telephone calls, faxes and emails to and from Ireland was in breach of the right to privacy, since relevant domestic law did not indicate with sufficient clarity the scope or manner in which to intercept and examine external communications and did not foresee adequate legal protection against the abuse of power. The law has to organize profiling and data mining powers in such a way that citizens *have an understanding* of the procedure to be followed when selecting for examination, sharing, storing and destroying intercepted material.

Profiling therefore needs a legal basis and legal safeguards need to be provided for the different stages of the process.[11]

Second, the interference has to be for a legal aim, which is relatively easy to fulfil given the broad range of aims listed in article 8(2).

Finally, the interference has to be 'necessary in a democratic society'. The notion of necessity does not involve a strict test of necessity in the sense of 'indispensable'. Nor is it as flexible as 'reasonable' or 'useful'. While it allows for a certain margin of appreciation, which can vary depending on the subject-matter, the scope for interpretation is limited.[12] The interference has to be a response to 'a pressing social need' and has to be 'proportionate to the legitimate aim pursued' (*Leander v. Sweden, Silver v. United Kingdom, Handyside v. United Kingdom*). The proportionality test used to give flesh to the bone of the necessity requirement involves four steps. The measure must be: (1) put in place to ensure a legitimate objective; (2) suitable, i.e. in a causal relation with the policy objective; (3) necessary, i.e. not curtailing rights more than is necessary given alternative options;[13] and (4) proportionate in a strict sense, i.e. even in the absence of a valid alternative, the benefits must outweigh the costs incurred by the infringement of the right.

This framework, including the proportionality test, said to be rooted in German administrative and constitutional law, is well anchored in the working praxis of both European Courts and national courts. But careful analysis of the relevant case law shows that both European and national courts often take a deferential approach towards public security or safety cases in order to leave some discretion to authorities and legislators, and therefore do not always fully assess the third and fourth steps of proportionality, namely the necessity and strict proportionality tests. Judges refrain from assessing choices made by governmental officials. Technically speaking, this is made possible by the lack of any explicit duty to apply the full proportionality test. Nowhere in case law do judges bind themselves to such strict testing. At the European level, the ECtHR in some cases suggests that not choosing the least onerous measure does not necessarily entail a violation of the ECHR or, more bluntly, explicitly rejects the test (Popelier and Van De Heyning 2013; Galetta and De Hert 2014).

Applying the framework: EUCJ data retention (2014) and German Rasterfahndung (2006)

The literature is mostly positive about the privacy case law of the two European Courts. Although the European Convention on Human Rights only mentions the right to privacy and does not mention modern technologies or the need to protect personal data, the ECtHR has opened up the Convention by incorporating most data protection safeguards in its case law. The mere storage of data often triggers the right to privacy; in the ensuing proportionality test, the specific aims of data collection are taken into account. A positive proportionality check for data collection is not automatically carried over to the retention or further use of the data (*S. and Marper v. United Kingdom*), which require separate proportionality considerations. The ECtHR does not oppose using databases containing data of earlier

offenders to investigate future crimes, but the persons affected must be limited and selected according to relevant criteria (*Van der Velden v. the Netherlands*). Retaining data because more data makes the system more useful is considered disproportionate (*S. and Marper v. United Kingdom, M.K. v. France*).

We wish to illustrate the potential effect of the European privacy framework for the practice of profiling with two other court decisions. First is the data retention decision of the European Court of Justice (EUCJ) of 8 April 2014. Directive 2006/24/EC obliged telecommunications and internet service providers to retain traffic, location and related data to identify the user, but not the actual communications, for 6 to 24 months and to allow access to this data by law enforcement agencies for the purpose of investigation, detection and prosecution of serious crimes. The data is not centralised by law enforcement agencies, which can access the data retained by service providers for specific cases. The obligation to retain data is a clear example of the growing surveillance engendered by the preventive approach to security.

The EUCJ bases its analysis on Article 7 of the Charter of Fundamental Rights of the European Union (CFREU), which is almost identical to Article 8(1) ECHR and Article 8 of the Charter. It considers both the retention of data by service providers and the access by law enforcement agencies as distinct interferences with the right to privacy. It points out that retention and subsequent use without informing concerned users "is likely to generate in the minds of the persons concerned the feeling that their private lives are the subject of constant surveillance". The broad retention also entails "an interference with the fundamental rights of practically the entire European population". The EUCJ points out that no differentiation or limitation is made linked to the objective of fighting serious crime. The directive therefore affects persons for whom there is no evidence linking their conduct to serious crime. Similarly, there is no relationship between the retained data and threats to public security. The EUCJ further points to the lack of adequate rules on the access and further use of the data, including criteria to determine the limits of such access and use, procedural rules and conditions, or criteria to limit persons and the retention period to the strictly necessary. The EUCJ therefore declared the directive invalid (EUCJ, C-293/12 - Digital Rights Ireland and Seitlinger and others).

Objections to broad interference – affecting the privacy of people having no connection to the targeted threat or criminal behaviour – pose a particular problem for profiling, especially the creation of profiles and the search for those who fit the profile. The other objections clarify that precise rules on access, use and retention of data are needed in any practice of large-scale profiling.

Another relevant case is the *Rasterfahndung* (data screening) case decided by the German Constitutional Court on 4 April 2006 (ECLI:DE:BVerfG:2006:rs2006 0404.1bvr051802). This decision was induced by the complaint of a Moroccan student against a large-scale data mining operation searching for possible terrorist sleeper cells after the terrorist attacks of 11 September 2001. First a dataset was made of persons fulfilling a set of criteria, which included being of the Muslim faith. In the state of Nordrhein-Westfalen, this dataset included about 11,000 persons, distilled from a set of 5,200,000 people. This dataset was then compared to find suspicious matches with a range of other databases, which contained data concerning between 200,000 and 300,000 people. This screening operation was followed by other investigative measures, but in the end yielded no results. The description lets us suppose that the operation involved the querying of databases, but probably made no use of data mining algorithms.

The German Constitutional Court declared the screening operation unconstitutional. The Court pointed out that the distinct steps of the screening operation were all interferences with the informational self-determination of people against whom no suspicion was present. Such interference could only be made proportional in limited circumstances. More specifically, a concrete danger or threat must be present and this determination must have factual grounds. General assessments of threat are insufficient, as this would lead to unrestricted competence and searches '*ins Blaue hinein*' (fishing expeditions).

These two privacy judgements set clear limits to the use of automated and predictive profiling. In general one can say that the ECtHR has transposed its guidelines on the establishment of safeguards and the minimum safeguards developed in its jurisprudence on secret listening to more modern practices such as profiling.

These practices need to serve a legitimate aim, need to be regulated with enough detail in hard law, and need to meet the proportionality test. Our two examples clarify that activities that interfere without distinction with the privacy of large parts of the population are disproportional and that such activities must in some way be linked to factual grounds. One has to admit, however, that the relevant case law is either very young or scarce, and does not allow stronger or more precise conclusions. Continued scrutiny is thus warranted. Investigative methods evolve together with technological possibilities; old standards can become outdated, even when laid down in the case law of our highest courts. Profiling and Big Data give the police 'something to work with', which might make their actions less discretionary and acceptable in light of existing human rights standards.

Ferguson, for instance, analyses the effects of the growing availability of data on the Fourth Amendment requirement of reasonable suspicion for stop and search interventions. The original 'small data' standard concerned the observable actions

of unknown subjects. The facts leading to suspicion had to relate to criminal activity, and not just the person. The availability of more data generally leads to more knowledge about an identified suspect and a resulting prediction feeding into the suspicion. The suspicion becomes less related to facts on actual activity and more related to the person. A suspicion based solely on the person would then allow stopping and searching that individual at any moment without further justification; a link to actual activity thus remains necessary to inform a suspicion (Ferguson 2014). This analysis can be generalized to practices of profiling to guide considerations of proportionality.

The decisions mentioned above show that requiring limitations based on factual grounds is a reasonable approach. The nuanced application of proportionality in the context of profiling requires further development, one which will maintain effective safeguards against the generalised application of intrusive methods as well as methods based on arbitrary or unfounded assumptions.

6.5 PROFILING AND THE PROHIBITION OF DISCRIMINATION

Article 14 ECHR and the 12th Protocol prohibit 'discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status' in the enjoyment of any right set forth in the Convention or by law. EU law contains similar non-discrimination principles.

The ECtHR considers discrimination to be "differences in treatment based on an identifiable characteristic, or 'status'" and checks for differences in the treatment of persons "in analogous, or relevantly similar, situations." The difference in treatment is discriminatory if it has "no objective and reasonable justification", if it does not pursue a "legitimate aim" or if there is no "reasonable relationship of proportionality between the means employed and the aim sought to be realised" (*Carson and others v. United Kingdom*). Generally, the principle of non-discrimination requires that comparable situations must not be treated differently and that different situations must not be treated in the same way (EUCJ, *Heinz Huber v. Germany*).

Data mining algorithms search for correlations in data; the resulting profile can make all sorts of characteristics a relevant difference. The non-discrimination principle limits what can be a relevant difference. The protected grounds are characteristics that should not be considered relevant, unless they can be adequately justified. This applies at different stages of the profiling process, from the selection of data to be used in profiling to the application of the profile.

Huber, delivered by the EUCJ, is an innovative judgement on the applicability of the non-discrimination principle in profiling practices.[14] The object of discussion was a database containing the personal details of foreigners applying for residence and its further use by police to fight crime.

The EUCJ concluded, firstly, that this register was not contrary to the law as far as it served the application of the residence legislation, contained only those personal data necessary for that application and granted access to other services which had competences in that field. A comparison with the registration of personal data of nationals in local registers with less personal data was made and the differences were accepted as justified as far as the centralised nature of the register allowed a more effective application of the residence law and it contained only data necessary for that purpose. In other words, the difference in treatment of personal data of foreigners compared with nationals was considered justified as it served legitimate purposes (application of residence legislation) and the difference in treatment was proportionate with that aim. The difference in treatment was objectively reasonably linked with a difference in legal situation, and therefore justified.

The same reasoning led to a different conclusion concerning the use of this register for crime fighting purposes. As the fight against crime involves the prosecution of crimes and offences committed irrespective of the nationality of their perpetrators, this objective cannot justify a difference in treatment between nationals and other EU citizens resident in the member state. A difference in treatment through the processing of personal data with a specific system for foreigners was therefore not justified for this purpose. Again, a comparison was made between foreign EU-citizens and nationals but here the situation of the two groups was too similar to justify a difference in treatment, although the aim was legitimate.

Huber underlines the relevance of the non-discrimination principle for profiling. While the case focused on the government's storage of and access to personal data, all other steps of the profiling process (see section 6.1) must be checked on non-discrimination grounds as well.

Firstly, there is the issue of linking data directly or indirectly to one of the protected grounds when designing profiles. The *Rasterfahndung* decision concerned a profile that included being of the Muslim faith. While the Court approached the case in terms of privacy, and did not make a separate evaluation of the non-discrimination principle, it did point to the risk of stigmatisation. Its general conclusion that such operations need a concrete threat sustained with factual elements also points to the obligation to justify the use of such sensitive criteria.

Secondly, there is a need to apply a non-discrimination test when assessing the results of the profiling process.[15] The application of the profile can result in indirect discrimination, as the 'neutral' application of algorithms leads to an unjustified burden on specific groups. This requires the application of the non-discrimination principle to the results of profiling and active checking of whether they lead to differential treatment that cannot be justified. Differential treatment is not excluded by definition. For example, suspects of terrorism inspired by religious, ethnic or ideological grounds will predominantly be of that religious, ethnic or ideological background, while suspects of social fraud will be recipients of social benefits. Nevertheless, differential treatment requires justification, and we can remind ourselves again of the requirement for a concrete threat sustained by factual grounds set by the German Constitutional Court.

These are just several examples of how profiles and data mining can violate the non-discrimination principle.

In their detailed study of the data mining and profiling process, Barocas and Selbst (2016) found no less than five discriminating mechanisms present at all steps in the process.[16] They include specifying the problem to be solved in ways that affect classes differently, failing to recognize or address statistical biases, reproducing past prejudice, and considering an insufficiently rich set of factors. Even in situations where data miners are extremely careful, there can still be discrimination when using models that, quite unintentionally, pick out proxy variables for protected classes. An additional problem is 'masking', when data miners are able to disguise intentional discrimination as unintentional.

With the exception of such masking, discriminatory data mining, Barocas and Selbst hold, is always unintentional. Evidence in court will be hard to produce, especially since discrimination cannot be legally blocked when there is a 'business necessity' (a US term), a pressing need or a reasonable justification. While both authors propose a range of non-legal solutions (oversampling, making training data and models auditable, pre-screening audits, results-focused balancing), they conclude that law will seldom work and that the market will not inspire costly efforts to do the profiling right, as a lot of Big Data profiling simply becomes ineffective when an absolute interdiction on protected classes or proxy variables is introduced. Other authors are less negative about the effects of anti-discrimination and see a role for specific auditing techniques (to be developed) to uncover hidden biases and for other technological solutions. For instance, the prohibition to discriminate on certain grounds can be modelled into discrimination-aware algorithms (Custers et al. 2013). The legal safeguard must be supported in this area by the development of technical safeguards and their implementation through standards and technical regulations. We return to this below.

Other human rights are affected by profiling, but the right to privacy and the prohibition of discrimination provide the basic safeguards for the problems we identified in our overview of risks. The procedural safeguards provided in Article 6 ECHR will be reviewed in our discussion of Dutch procedural safeguards that apply to profiling for investigative purposes, below.

6.6 ADDITIONAL GUARANTEES IN DUTCH ADMINISTRATIVE, CRIMINAL PROCEDURE AND DATA PROTECTION LAW?

The guarantees offered by the legality and purpose specification principle

While fundamental rights are an important set of checks on governmental powers, other safeguards elaborated in specific areas of law might be relevant as well. We now turn to existing legal safeguards in administrative, criminal procedure and data protection law.

A first safeguard can be found in the legality principle in administrative law: this principle gives flesh to the idea of the rule of law and the need to limit state powers by linking them to competences and purposes. Public authorities and their officials are only allowed to impact on citizens' freedoms when they have a competence to do so provided by law. Public authorities and officials have different roles and are therefore also provided with different competences. Linked to the legality principle is the speciality principle: a law can only be applied in its specific domain and therefore competences can only be used for the purposes provided by this law.

Dutch administrative law embeds this in the prohibition of *détournement de pouvoir* or the prohibition to use competences for other purposes but those for which they were provided. Similarly, in criminal procedure law investigation measures need to be based on law. This requirement not only results from the fundamental rights provisions in the ECHR (see our discussion of the legality requirement in Article 8 ECHR above) but also from 'the principle of procedural legality' found in Article 1 of the Code of Criminal Procedure (CCP) of 1926. Data protection law has its specific application of the legality principle through the purpose limitation principle (personal data may only be collected for specific, explicitly defined and legitimate purposes) and the requirement for legal grounds to process data.

Closer examination reveals that these principles have grey zones that are skilfully used by authorities to operate without detailed legislation. The principle of procedural legality (Article 1 CCP), for example, is interpreted in a restrictive way as covering only those methods of investigation that substantially infringe on fundamental rights. Investigative methods that do not, or do not substantially, breach fundamental rights can always be employed. Article 2 of the Police Act (*Politiewet*) on the task of the police is considered by the courts to provide a sound basis for investigative methods (Van Kempen 2009).

Equally vague is the *Wet bescherming persoonsgegevens* (Wbp), the Dutch Data Protection Act implementing directive 95/46/EC. Article 8 allows the processing of data on the basis of consent (Article 8a) or in order to comply with legal obligations (Article 8c), but then opens the spectre by stating that data processing is also possible when "processing is necessary for the proper performance of a public law duty by the administrative body concerned or by the administrative body to which the data are provided" (Article 8d). Article 9 §1 Wbp adds that "personal data shall not be further processed in a way incompatible with the purposes for which they have been obtained". In other words, in limited circumstances, further processing for another purpose remains possible. These provisions raise important questions about the resilience of the Dutch Data Protection Act (and other similar acts in EU Member States that have identical provisions). Note that Article 9§2 contains provisions to limit the possible abuse of the 'compatibility' clause.[17] The Act on Police Data (*Wet politiegegevens* (Wpg)) contains a similar purpose limitation principle and the grounds for processing, including a specific ground for automated comparison and the combined searching of data.

To the extent that authorities interpret the vagueness in criminal and data protection law to allow profiling, the intensive exploitation of existing data through the use of profiles and data mining will go unnoticed by lawmakers and citizens, since these are considered to be implicit ('compatible') or not 'substantially infringing' powers. This development will leave the automated administrative state largely unchecked by legal rules.

There are signs that this development is already well under way. In the memorandum accompanying the SUWI proposal (see above), the government explained its choice for a broad formulation of purpose in order to allow a range of public authorities to cooperate and act as an integrated public authority. It is thereby silently assumed that this purpose is compatible with the original purpose for which data were collected by participating data providers. SyRI is a framework in which a different group of partners more specifically define the content and purpose of data exchange for each project; this limits the loosening of the principle. SyRI has now become the model for developing a more general legal framework for the exchange of data between public and possibly private partners (Werkgroep Verkenning kaderwet gegevensuitwisseling 2014). In the context of Big Data and the Internet of Things, Moerel and Prins have recently advocated exchanging the purpose limitation principle for a legitimate purpose principle; they point to recent proposals by the European Council to allow further processing even for an incompatible purpose when the legitimate interests of the controller or a third party override the interests of the data subject (Moerel and Prins 2015). In fact we see a similar shift towards a legitimate interest principle in these proposals for cooperation structures, especially when the purpose of data processing is as broadly formulated as the public task of the authorities involved.

Guarantees offered by the proportionality, subsidiarity and necessity principles

A second safeguard is the proportionality principle present in all three legal frameworks. In administrative law the proportionality principle (*evenredigheid*) states that decisions may not affect persons disproportionately, compared to the purpose of the decision. Supervisory and investigative competences may only be used when needed. In criminal procedure law, the proportionality principle is embedded in the conditions, such as the presence of reasonable suspicion, that must be fulfilled before certain investigative measures can be taken. The Dutch Data Protection Act states that personal data shall only be processed when adequate, relevant and not excessive (Article 11). The subsidiarity principle (the purpose cannot be reached with less negative impact) and the necessity principle (the interference is needed to attain the purpose) are linked to the proportionality principle.

In our discussion of the right to privacy (see above), we observed that European and national courts have held back in testing necessity and strict proportionality, thereby allowing some discretion to authorities. Here we focus on the requirement in criminal law that investigative powers can only be used in case of reasonable suspicion that an offence (infraction or crime) has been committed (cf. Article 27 CCP). This is an important safeguard as it prevents, by using a threshold requirement, wide use of the state's far-reaching powers in criminal law. But the rise of supervisory powers other than investigative powers and developments that lower the threshold are thwarting this safeguard. These will be discussed below.

An important distinction in Dutch administrative law is that between supervisory powers (*toezicht*) and investigative powers (*opsporing*). Supervisory powers do not aim to investigate offences but to observe the application of regulations. The use of these powers is not linked to any suspicion. As part of these supervisory powers, officials can demand information or documents. Citizens are obliged to cooperate; refusing to do so is an offence. Supervisory powers are generally part of administrative law and have a proactive or preventive role. In contrast, investigative powers aim to investigate offences with prosecution as their objective. They are used to respond to an offence, to establish the facts and the guilt of the offender. In the traditional view they are linked with a suspicion of guilt for an offence. Investigative powers can be found in both criminal procedure and administrative law, and are exercised under the control of the public prosecutor.

An intermediary position is taken by powers of (repressive) control. These powers are similar to supervisory powers in that they are used without the presence of a suspicion, but with the objective of uncovering offences (e.g. traffic controls). These powers are traditionally included within the legal framework on investigative powers (Borgers 2011).

The distinction between supervisory and investigative powers is important in light of the prohibition of *détournement de pouvoir* as it raises the question of how the subsequent, concurring or overlapping use of different competences is addressed. It turns out that the courts deal leniently with such issues. The use of investigative powers when supervisory powers have uncovered an offence is an obvious, non-problematic response. But questions arise when supervisory powers tied to a specific part of legislation uncover an offence in the ambit of another law. Another problem is the continued use of supervisory powers when investigative powers are employed, or their concurring use. Both have been accepted by the Supreme Court, in the latter case through the procedural rights of investigative measures. The exception is when a competence from one law is used exclusively to obtain the objectives of another law (Borgers 2011).

The link between investigative powers and the condition of reasonable suspicion has been loosened with the introduction of more proactive and preventive investigative measures. With the introduction of a terrorist crime, the condition of reasonable suspicion was lowered to 'indications' of a terrorist crime (Hirsch Ballin 2008; Van Kempen 2009). Article 126gg Sv. regulates the 'exploratory investigation' (*verkennend onderzoek*) that precedes the proper investigation, which considers the presence of crimes or their planning among groups of people. It consists of the collection and analysis of information from police databases or open sources and is not based on a suspicion against a specific person. Exploratory investigations begin based on 'indications following from facts or circumstances'. Mere presumptions are not enough; factual grounds are required. When the exploratory investigation concerns a terrorist crime, Article 126hh Sv. allows summoning the delivery of other databases, including from private data holders. This data can be compared or processed together with other datasets. Article 126gg Sv. still applies.

Articles 126gg and 126hh Sv. regulate data screening operations within the criminal procedure, which may include profiling. Their preliminary character implies that data screening operations remain outside the safeguards of Article 6 ECHR. The main safeguard is that these operations happen under the control of the public prosecutor.

We can conclude that there is a general trend towards allowing data exchange and profiling by loosening existing safeguards in administrative, criminal and data protection law. The *détournement de pouvoir* principle in administrative law was hollowed out much earlier. In criminal procedure we notice the emergence of less stringent conditions to use data-intensive investigative methods in a much earlier phase. In data protection we notice the diminished impact of the purpose limitation principle.

On the one hand, this is done for good reasons. Some of the safeguards were developed for different technological and societal circumstances. The strict application of *détournement de pouvoir* makes good sense for Weberian bureaucratic organisations, but less sense for networked and interconnected organisations sharing tasks. Although the strict application of purpose limitation was logical for isolated databases, the linking of data sources is crucial for the new data-intensive methods.

While successful prevention is clearly beneficial for society, it raises the question of whether attention is now too focussed on loosening safeguards rather than reinventing or adapting them. In light of the proportionality principle, the negative impacts of heightened surveillance are neglected.

Guarantees offered by procedural safeguards

Procedural safeguards for individuals involved in state procedures are a third category of safeguards. Dutch administrative law contains a duty of care. It obliges the administration to carefully establish and review all relevant factual and legal elements of a case. This includes checking whether advisors have properly carried out the research requested of them. To give them the opportunity to be heard, the administration in certain situations must inform affected parties when preparing a decision. When the authority bases its decision on a technical investigation, the result must be documented in a report to allow later review.

The procedural safeguards offered by Article 6 ECHR play a similar role in criminal procedural law. The procedural rights of a fair trial involve the right of access to a court, an independent and impartial tribunal established by law. It further involves the right to a fair hearing, which includes the right to an adversarial trial, freedom from self-incrimination, and so on. Criminal procedures have extra safeguards such as the assumption of innocence (Harris et al. 2009). Guarantees in criminal procedures apply from the moment a person is charged with a criminal offence; they do not apply to pre-trial investigations or preventive measures preceding a criminal charge. The impact of these rights can be seen in the duty to cooperate with supervisory powers. Once investigative measures are employed, the freedom from self-incrimination comes into play. This also applies to administrative sanctions when these have punitive or deterring objectives (Borgers 2011).

A range of similar safeguards are recognized in data protection law ('data subject rights'): the right to be informed, to access the data, to have it corrected or (when no longer relevant) deleted, and the right to object to processing. Nobody may be subjected to automated decision-making with legal consequences when such decisions are based only on data intended to provide a picture of certain aspects of their personality. Exceptions exist, in which case persons must be allowed to

present their views and be informed of the logic on which the automated decision is made. This concurs with the duty of care, which obliges a review of preparatory research and to allow affected persons to be heard.

The SyRI system is an example of how these safeguards are applied to profiling. While the system flags persons linked to suspicious data patterns, this flagging is only guidance for further investigation and does not lead to automatic decisions. Decisions are only taken after investigation and thereby subjected to the whole range of procedural safeguards. Risk notifications are included in a register. People can inquire if they are included in the register, but are not informed directly when they show up in a risk notification.

While profiling is subject to an extensive range of procedural safeguards, Citron warns that these safeguards are imperilled by automation. She points out that automation bias and the opacity of data processing can diminish or eliminate the distinction between automated and computer-assisted decision-making. Exaggerated trust in computers runs counter to the duty of care. EUCJ case C-503/03 between the European Commission and Spain, concerning the refusal of entry into the Schengen area based on flagging in the SIS system, confirms this duty of care. The EUCJ pointed out that such refusal without preliminary verification of whether the person presented an actual danger violated EU law.

Citron points to the need to more effectively implement safeguards. The opacity of automated processes imperils the flow of information to data subjects as well as the proper review of the basis of decisions (in the case of profiling, the flagging of a person as a risk). Proper audit trails, documenting the rules applied and the data considered, should be a required part of the system. Decision-makers should also explain in detail how they relied on computer-generated information, and be trained to critically evaluate such information to combat automation bias. Automated systems should be designed with transparency and accountability as core objectives. Citron advises that code be made available as open source; for profiling, this implies that algorithms and profiles are open to review. As code and algorithms can function as hidden rule-makers, this will provide an alternative means of external scrutiny. For the same reason, Citron advises allowing the public to participate in reviewing the systems. Lastly, she makes a plea for proper testing.

In the case of profiling, this advice can be further specified to require algorithms to be discrimination-aware and to develop auditing and testing protocols for algorithms and profiles (Pedreschi et al. 2013; Romei and Ruggieri 2013). Making algorithms and (risk) profiles open to public review should be a priority. If such transparency facilitates anticipation and avoiding behaviour, the internal auditing and publication of results still remain possible. Protocols to measure or assess the

control burden can also be developed and integrated within new technical standards for profiling. Impact assessments foreseen in the draft General Data Protection Regulation can be widened in scope to include aspects such as discrimination.

6.7 CONCLUSION: LEGAL SAFEGUARDS MUST BE OVERHAULED TO MAKE THEM EFFECTIVE AGAIN

We distinguished between the creation of a profile and its application, as they affect different groups of people. We clarified that automated profiling based on data mining is not yet by definition an application of Big Data, but that Big Data magnifies its impact. Profiling in a Big Data context puts greater pressure on the checks and balances in the legal framework.

We then discussed examples of profiling in the US and the Netherlands. Predictive policing, an application of profiling for security purposes, is widely used in the US and is seeing its first applications in the Netherlands. We also presented two Dutch examples of profiling to combat fraud: the iCOV and the SyRI system.

While profiling has clear benefits in guiding efforts and improving efficacy, it comes with several risks. We highlighted three potential risks: the intrusion on privacy, social sorting and discrimination, and opaque decision-making. We therefore reviewed the legal safeguards present in the human rights framework and in Dutch administrative, criminal procedure and data protection law and highlighted the relevant cases. Aware of the potential to discriminate and infringe on privacy, courts insist on proportionality, including clear and detailed rules and requirements such as reasonable indications or suspicious before making larger groups of people the object of governmental actions.

Our review also covered more critical concerns. If courts refrain from testing the necessity and strict proportionality of governmental data mining and use of profiles, these technologies may well go unchecked. If these practices are seen as 'lesser' infringements that do not require detailed regulation, the resilience of the legal framework might similarly be low. Given the opacity of profiles and data mining operations, Barocas and Selbst (2016) point to the difficulty of triggering legal guarantees as well as legal contradictions. There might, for instance, be free speech objections against prohibiting governments and other actors from using certain data when assembling or applying profiles. One paradox we noted concerned the issue of reasonable suspicion and similar thresholds. What will courts do when authorities invoke not facts, but the 'hits' and 'outcomes' of profiling operations to argue that an individual or group of individuals need extra surveillance or investigation? With Ferguson (2014), we believe that the growing availability of data might lead to an erosion of existing standards, rendering police and governmental interventions legitimate because the computer 'said so'.

Our review revealed tensions between the use of profiling as a proactive investigative technique and the legal safeguards in data protection, administrative and criminal law. These tensions have been resolved by loosening legal safeguards. The *détournement de pouvoir* principle in administrative law was hollowed out much earlier. In criminal procedure we see the emergence of less stringent conditions to use data-intensive investigative methods in a much earlier phase. In data protection we notice the diminished impact of the purpose limitation principle. In considerations of the proportionality principle, the negative impacts of heightened surveillance are neglected.

Human rights jurisprudence on the right to privacy sets clear limits to the use of automated and predictive profiling. It makes clear that profiling and data screening which interfere without distinction with the privacy of large parts of the population are disproportional and that such activities must in some way be linked to concrete and factual elements. On the other hand, the jurisprudence is scarce; drawing strong or precise conclusions remains difficult. While the prohibition of discrimination is a useful legal safeguard, it must be given teeth through the development of audit tools and discrimination-aware algorithms.

Although done for good reasons, the question is whether the focus has not been too much on loosening safeguards rather than adapting them. We concur with Citron that legal safeguards need to be overhauled to make them effective again in the 'automated administrative state'. Transparency and accountability can be designed into profiling practices. This must be backed up with stronger institutional safeguards to allow for independent assessments and rapid feedback into decision-making.

REFERENCES

Barocas, S. and A. Selbst (2016) 'Big Data's Disparate impact', *California Law Review* 104: 1-60 (forthcoming).

Black, J. (2005) 'The Emergence of Risk-Based Regulation and the New Public Risk Management in the United Kingdom', *Public Law* Autumn: 512-549.

Borgers, M.J. (2011) 'De onderzoeksfase: toezicht, controle en opsporing', pp. 455-496 in F.G.H. Kristen, R.M.I. Lamp, J.M.W. Lindeman and M.J.J.P. Luchtman (eds.) *Bijzonder strafrecht. Strafrechtelijke handhaving van sociaal-economisch en fiscaal recht in Nederland*, The Hague: Boom Lemma.

Brakel, R. van and P. De Hert (2011) 'Policing, Surveillance and Law in a Pre-Crime Society: Understanding the Consequences of Technology Based Strategies', *Technology-Led Policing* 20: 165.

Chicago Police Department (2013) *Custom Notifications in Chicago – Pilot Program D13-09*, available at: http://directives.chicagopolice.org/directives-mobile/data/a7a57bf0-13fa59ed-26113-fa63-2e1d9a10bb60b9ae.html?ownapi=1.

Citron, D.K. (2008) 'Technological Due Process', *Washington University Law Review* 85: 1249-1313.

College Bescherming Persoonsgegevens (2014) *Advies conceptbesluit SyRI*, available at: https://cbpweb.nl/sites/default/files/atoms/files/z2013-00969.pdf.

Custers, B. (2014) 'Risicogericht toezicht, Profiling and Big Data', *Tijdschrift voor Toezicht* 5, 3: 9-16.

Custers, B., T. Calders, T. Zarsky and B. Schermer (2013) 'The Way Forward' in B. Custers, T. Calders, B. Schermer and T. Zarsky (eds.) *Discrimination and Privacy in the Information Society*: 341-57, Berlin/Heidelberg: Springer.

Fayyad, U., G. Piatetsky-Shapiro and P. Smyth (1996) 'From Data Mining to Knowledge Discovery in Databases', *AI Magazine* 17, 3: 37-54.

Federal Trade Commission (2014) *Data Brokers: a Call for Transparency and Accountability*, available at: www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade- commission-may-2014/140527databrokerreport.pdf.

Ferguson, A.G. (2014) 'Big Data and Predictive Reasonable Suspicion', *University of Pennsylvania Law Review* 163: 327.

Galetta, A. and P. De Hert (2014) 'Complementing the Surveillance Law Principles of the ECtHR with Its Environmental Law Principles: An Integrated Technology Approach to a Human Rights Framework for Surveillance', *Utrecht Law Review* 10, 1: 55-75.

Gandy Jr., O.H. (2010) 'Engaging Rational Discrimination: Exploring Reasons for Placing Regulatory Constraints on Decision Support Systems', *Ethics and Information Technology* 12, 1: 29-42.

Harcourt, B.E. (2005) 'Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age', *University of Chicago Public Law Working Paper* 94.

REFERENCES

Barocas, S. and A. Selbst (2016) 'Big Data's Disparate impact', *California Law Review* 104: 1-60 (forthcoming).

Black, J. (2005) 'The Emergence of Risk-Based Regulation and the New Public Risk Management in the United Kingdom', *Public Law* Autumn: 512-549.

Borgers, M.J. (2011) 'De onderzoeksfase: toezicht, controle en opsporing', pp. 455-496 in F.G.H. Kristen, R.M.I. Lamp, J.M.W. Lindeman and M.J.J.P. Luchtman (eds.) *Bijzonder strafrecht. Strafrechtelijke handhaving van sociaal-economisch en fiscaal recht in Nederland*, The Hague: Boom Lemma.

Brakel, R. van and P. De Hert (2011) 'Policing, Surveillance and Law in a Pre-Crime Society: Understanding the Consequences of Technology Based Strategies', *Technology-Led Policing* 20: 165.

Chicago Police Department (2013) *Custom Notifications in Chicago – Pilot Program* D13-09, available at: http://directives.chicagopolice.org/directives-mobile/data/a7a57bf0-13fa59ed-26113-fa63-2e1d9a10bb60b9ae.html?ownapi=1.

Citron, D.K. (2008) 'Technological Due Process', *Washington University Law Review* 85: 1249-1313.

College Bescherming Persoonsgegevens (2014) *Advies conceptbesluit SyRI*, available at: https://cbpweb.nl/sites/default/files/atoms/files/z2013-00969.pdf.

Custers, B. (2014) 'Risicogericht toezicht, Profiling and Big Data', *Tijdschrift voor Toezicht* 5, 3: 9-16.

Custers, B., T. Calders, T. Zarsky and B. Schermer (2013) 'The Way Forward' in B. Custers, T. Calders, B. Schermer and T. Zarsky (eds.) *Discrimination and Privacy in the Information Society*: 341-57, Berlin/Heidelberg: Springer.

Fayyad, U., G. Piatetsky-Shapiro and P. Smyth (1996) 'From Data Mining to Knowledge Discovery in Databases', *AI Magazine* 17, 3: 37-54.

Federal Trade Commission (2014) *Data Brokers: a Call for Transparency and Accountability*, available at: www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade- commission-may-2014/140527databrokerreport.pdf.

Ferguson, A.G. (2014) 'Big Data and Predictive Reasonable Suspicion', *University of Pennsylvania Law Review* 163: 327.

Galetta, A. and P. De Hert (2014) 'Complementing the Surveillance Law Principles of the ECtHR with Its Environmental Law Principles: An Integrated Technology Approach to a Human Rights Framework for Surveillance', *Utrecht Law Review* 10, 1: 55-75.

Gandy Jr., O.H. (2010) 'Engaging Rational Discrimination: Exploring Reasons for Placing Regulatory Constraints on Decision Support Systems', *Ethics and Information Technology* 12, 1: 29-42.

Harcourt, B.E. (2005) 'Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age', *University of Chicago Public Law Working Paper* 94.

Harris, D., M. O'Boyle, E. Bates and C. Buckley (2009) *Harris, O'Boyle and Warbrick: Law of the European Convention on Human Rights*, Oxford: Oxford University Press.

Hildebrandt, M. (2008a) 'Defining Profiling: A New Type of Knowledge?' in M. Hildebrandt and S. Gutwirth (eds.) *Profiling the European Citizen*, Dordrecht: Springer Science + Business Media.

Hildebrandt, M. (2008b) 'Profiling and the Rule of Law', *Identity in the Information Society* 1, 1: 55-70.

Hirsch Ballin, M.F.H. (2008) 'Inside View of Dutch Counterterrorism Strategy: Countering Terrorism through Criminal Law and the Presumption of Innocence', *Journal of the Institute of Justice and International Studies* 8: 139-51.

Kempen, P.H. van (2009) 'The Protection of Human Rights in Criminal Law Procedure in The Netherlands', *Electronic Journal of Comparative Law* 13, 2: 37.

Kitchin, R. (2014a) *The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences*, London: Sage.

Kitchin, R. (2014b) 'Big Data, New Epistemologies and Paradigm Shifts', *Big Data and Society* 1, 1 DOI: 10.1177/2053951714528481.

McCaney, K. (2013) 'Prisons Turn to Analytics Software for Parole Decisions', GCN, available at: http://gcn.com/articles/2013/11/01/prison-analytics-software.aspx.

Moerel, L. and C. Prins (2015) *Further Processing of Data Based on the Legitimate Interest Ground: The End of Purpose Limitation?*, Tilburg: Tilburg University.

Openbaar Ministerie (2014) *Aanwijzing opsporingsbevoegdheden (2014A009)*, available at: www.om.nl/organisatie/beleidsregels/overzicht-0/opsporing-politie/@86281/aanwijzing-3/.

Pedreschi, D., S. Ruggieri and F. Turini (2013) 'The Discovery of Discrimination', pp. 91-108 in B. Custers, T. Calders, B. Schermer and T. Zarsky (eds.) *Discrimination and Privacy in the Information Society*, Berlin/Heidelberg: Springer.

Perry, W.L. et al. (2013) *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*, Rand Corporation.

Popelier, P. and C. van de Heyning (2013) 'Procedural Rationality: Giving Teeth to the Proportionality Analysis', *European Constitutional Law Review* 9: 230-262.

Romei, A. and S. Ruggieri (2013) 'Discrimination Data Analysis: A Multi-Disciplinary Bibliography, pp. 109-135 in B. Custers, T. Calders, B. Schermer and T. Zarsky (eds.) *Discrimination and Privacy in the Information Society*, Berlin/Heidelberg: Springer.

Schermer, B. (2013) 'Risks of Profiling and the Limits of Data Protection Law, pp. 137-152 in B. Custers, T. Calders, B. Schermer and T. Zarsky (eds.) *Discrimination and Privacy in the Information Society*, Berlin/Heidelberg: Springer.

Senate Committee on Commerce, Science and Transportation (2013) *A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes*, available at: www.commerce.senate.gov/public/?a=Files.Serve&File_id=0d2b3642-6221-4888-a631-08f2f255b577.

Solove, D.J. (2004) *The Digital Person: Technology and Privacy in the Information Age*, New York: NYU Press.

United States Government Accountability Office (2004) *Data Mining: Federal Efforts Cover a Wide Range of Uses*, GAO-04-548, available at: www.gao.gov/new.items/d04548.pdf.

United States Government Accountability Office (2013) *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, GAO-13-663, available at: www.gao.gov/assets/660/658151.pdf

Vedder, A. (1999) 'KDD: The Challenge to Individualism', *Ethics and Information Technology* 1, 4: 275-281.

The Verge (2014) *The Minority Report: Chicago's New Police Computer Predicts Crimes, But is it Racist?*, available at: www.theverge.com/2014/2/19/5419854/the-minority-report-this-computer-predicts-crime-but-is-it-racist.

Werkgroep Verkenning kaderwet gegevensuitwisseling (2014) *Kennis delen geeft kracht: Naar een betere èn zorgvuldigere gegevensuitwisseling in samenwerkingsverbanden*, available at: http://njb.nl/Uploads/2015/1/blg-442395.pdf.

NOTES

1 An exhaustive analysis of the legal framework within which profiling takes place is not possible in this limited space, but our main question is if this legal framework provides adequate safeguards against the risks mentioned above. The first stage of this analysis is the human rights framework. We then consider Dutch administrative law and criminal procedure.

2 Profiling by police or border control officers during security checks when based on a range of behavioural elements provides a more legitimate example of old fashioned human profiling.

3 Profiling is used in this way in our examples thus far, which is also its main actual use.

4 But Citron (2008) shows how the distinction between automated decision-making and profiling as a decision support tool can become superficial in practice. In our analysis of legal safeguards below, we focus on these uses as they entail the most risk of negative impacts.

5 We implicitly assume that the data and profiling concern persons, but this is not necessarily the case. Profiling can also concern objects, places or other phenomena. For example, profiling can be used in industrial processes to identify defective components or, in the context of customs, to discern ships or containers with a higher risk of illegal imports or places and times with a higher risk for criminal activities.

6 These methods forecast places and times with an increased risk of crime.

7 These methods identify individuals at risk of offending in the future.

8 These techniques are used to create profiles that accurately match likely offenders with specific past crimes.

9 These approaches are used to identify groups or, in some cases, individuals who are likely to become victims of crime.

10 A similar system is in use in: the 'Criminaliteits Anticipatie Systeem' (CAS).

11 But in the Court's reasoning, the expression 'in accordance with law' is compatible with the establishment of national legal regimes to regulate differential privacy interferences. When regulating 'soft' privacy interferences, national law is given greater flexibility and a wider margin of appreciation for remedies used to counter the concerned interference and its negative effects. It is not apparent from the reasoning of the Court how to classify those interferences that imply both monitoring and tracking such as profiling, data mining and Internet monitoring in general. At present, this is an open question in European case law (Galetta and De Hert 2014).

12 The margin of appreciation of the state also depends on factors like the nature and seriousness of the interests at stake and of the interference (*Peck v. United Kingdom*, *Z. v. Finland*, *Leander v. Sweden*).

13 A subsidiarity check – whether the legitimate aim could not be obtained through less intrusive and therefore more proportionate means – is found in *Peck v. United Kingdom*.

14 Although it concerned a EU citizen and was based on the prohibition of discrimination on the ground of nationality in EU law, it applied a similar reasoning as the ECtHR.

15 Data mining is sometimes presented as a guarantee against discrimination because its algorithms allow a more objective treatment by excluding human bias (Custers 2014). This view is too optimistic and forgets that such human bias can seep through at all stages of the profil-

ing process, beginning with the selection of data. Data mining algorithms will rather objectively reflect in their results the presence of such human bias in the original data. Therefore one of the main dangers linked with profiling is to make discrimination hidden and as opaque as the functioning of data mining algorithms.

16 Their description is slightly more complex than ours in section 6.1, and distinguishes between: defining 'the target variable', labeling and collecting the 'training data', 'feature selection', and making decisions on the basis of the resulting model.

17 "For the purposes of assessing whether processing is incompatible, as referred to under (1), the responsible party shall in any case take account of the following:

a. the relationship between the purpose of the intended processing and the purpose for which the data have been obtained;
b. the nature of the data concerned;
c. the consequences of the intended processing for the data subject;
d. the manner in which the data have been obtained, and
e. the extent to which appropriate guarantees have been put in place with respect to the data subject".